中国科技核心期刊
ISSN 1009-671X
CN 23-1191/U

数据中心IPv6网络日志采集平台的设计与实现

吴宇平, 邢丽刃, 秦绪杰

引用本文:
吴宇平,邢丽刃,秦绪杰. 数据中心IPv6网络日志采集平台的设计与实现[J]. 应用科技, 2022, 49(3): 76–83.
WU Yuping,XING Liren,QIN Xujie. Design and implementation of IPv6 network log collection platform in data center[J]. *Applied science and technology*, 2022, 49(3): 76–83.

在线阅读 View online: https://dx.doi.org/10.11991/yykj.202112007

DOI: 10.11991/yykj.202112007
网络出版地址: https://kns.cnki.net/kcms/detail/23.1191.U.20220316.1504.002.html

数据中心 IPv6 网络日志采集平台的设计与实现

吴宇平，邢丽刃，秦绪杰

哈尔滨工程大学 信息化处，黑龙江 哈尔滨 150001

摘　要: 为在高校数据中心 IPv6 建设工程中，有效进行 IPv6 的管理与运维，本文构建了 IPv6 网络日志采集平台。通过对 IPv6 网络日志采集平台进行研究与设计，采用大数据技术进行处理与分析，构建支持 IPv6 网络管理运维的可视化的监控体系，实现高校数据中心 IPv6 管理与运维效率的提高。

关键词: IPv6; 数据中心; 日志采集; 日志分析; 网络运维; 网络安全; 大数据; 可视化监控
中图分类号: TP393.4　　**文献标志码:** A　　**文章编号:** 1009−671X(2022)03−0076−08

Design and implementation of IPv6 network log collection platform in data center

WU Yuping , XING Liren, QIN Xujie

Informationization Office, Harbin Engineering University, Harbin 150001, China

Abstract: This paper mainly explicates the construction of an IPv6 network log collection platform in order to effectively carry out IPv6 management, operation and maintenance in the IPv6 construction project of university data centers. Through the research and design of the IPv6 network log collection platform, the big data technology is used for data processing and analysis, and a visualized monitoring system that supports IPv6 network management, operation and maintenance is constructed, which could improve the efficiency of IPv6 management, operation and maintenance of university data centers.

Keywords: IPv6; data center; log collection; log analysis; network operation and maintenance; network security; big data; visual monitoring

近几年，伴随《推进互联网协议第六版（IPv6）规模部署行动计划》、《教育信息化 2.0》以及《教育现代化 2035》一系列重磅文件颁布，IPv6 在国内高校范围内获得大面积重点部署。截止 2021 年 11 月，我国 IPv6 互联网活跃用户达到 5.685 亿，占比约 56%，其中双一流大学网站 IPv6 占比更是达到了近 85%[1]，高校数据中心 IPv6 网络的管理与运维等问题也随之变为热点与难点，如基于 IPv6 的攻击已经开始出现，分布式拒绝服务攻击、网络穿透攻击、IPv6 加密蠕虫等开始出现于安全媒体，但却没有引起安全界的重视[2]。

通过建设 IPv6 网络日志采集平台，采集高校数据中心 IPv6 网络流量日志、域名系统（domain name system, DNS）日志以及相关网络设备、安全设备等日志，对 IPv6 地址、端口、应用协议等日志数据进行可视化监控、查询，构建一套以 IPv6 为研究核心的网络管理运维的可视化的监控体系。在网络安全管理方面，通过日志平台综合分析，能及时了解校园网网络状态、攻击来源、受攻击情况，建立威胁可视化及分析能力，检测出潜在的、恶意的攻击行为[3]，实现高校数据中心 IPv6 管理与运维工作的规范化，为高校更好地建设、管理与应用数据中心 IPv6 网络提供数据依据。

1　IPv6 网络日志采集平台构建

IPv6 网络日志采集平台主要采用 Cloudera 公司的 CDH-Hadoop 作为核心基础搭建。主要功能是对 IPv6 网络日志进行数据采集、解析、存储及大数据分析计算，并对日志进行相应的操作和统计，把检测 IPv6 数据包产生的报警和日志以各种

收稿日期: 2021−12−09.　　**网络出版日期:** 2022−03−17.
基金项目: 赛尔网络下一代互联网技术创新项目（NGII20190115）.
作者简介: 吴宇平, 男, 高级工程师, 博士.
通信作者: 吴宇平, E-mail: wyp@hrbeu.edu.cn.

丰富的格式进行记录，以方便查询和分析，可以为网络分析提供依据[4]。

1.1　CDH-Hadoop 概述

Hadoop 是一种开源的分布式编程软件框架，同时还是可以开发运行分布式应用程序、处理大数据的平台，它的 2 个核心组成部分为 HDFS 和 Map Reduce[5]。

Hadoop 本身是分布式架构，而且本身包括的管理组件数量很多，因而相应的部署和运维的门槛较高，所以市场出现了很多不同公司基于 Hadoop 再开发的产品。部署较多的有 Cloudera 公司的 CDH、华为的 MRS 云服务、新华三的 H3C DataEngine 大数据平台等发行版本。

1.2　平台设计原则

IPv6 集成了 IPsec 协议，基于 IPsec 协议可实现 IPv6 网络的加密通信，但请求注解 (request for comments, RFC）规定默认关闭 IPsec 协议，从而保证了 IPv6 网络日志采集的数据为非加密数据。本文设计的数据中心 IPv6 网络日志采集平台主要是采集各类网络设备、系统本身记录存储的 IPv6 网络相关日志，并进行存储、解析、查询与分析。

高校数据中心 IPv6 网络日志包括防火墙日志、DNS 日志、网站应用级入侵防御系统（web application firewall,WAF）日志以及深度报文检测 (deep packet inspection,DPI) 设备日志等多种数据源日志，数据量大，且不同厂商日志数据节点分散、格式不统一。从多源 IPv6 网络日志的特征考虑，平台设计需要满足系统的可扩展性、兼容性以及高效性。

平台进行数据采集设计时，要拥有通用且集成接口，以确保可以兼容各类网络以及安全设备，保障平台未来发展的可扩展性。

平台进行数据解析设计时，应着重考量具备不同的取值方式、兼容不同类型的数据源、过滤冗余数据、提取关键信息的能力。

平台进行数据分析设计时，由于多源 IPv6 网络日志数据量大，产生的周期短，因此系统需要具备针对海量数据的高速分析处理能力[6]。

1.3　平台整体架构

IPv6 网络日志采集平台由数据采集平台和数据分析平台两部分组成。数据采集平台为数据处理部分，采用 Cloudera 商用 Hadoop 集群搭建[7]；数据分析平台为应用部分，利用 B/S 架构进行设计开发，采取可编辑式的可视化界面实现数据可视化展现及监控。具体架构如图 1 所示。

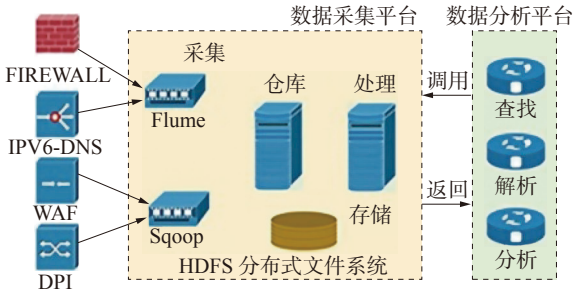

图1　IPv6 网络日志采集平台架构

1.4　平台工作原理

将数据中心核心交换机流量日志数据、DNS 解析日志数据、防火墙流量日志数据的 syslog 通过用户数据报协议 (user datagram protocol, UDP) 的方式发送到数据采集平台的 Flume 组件，进行协议解析并将数据以副本的方式存储到数据采集平台的 HDFS 分布式文件系统上；WAF日志数据以及 DPI 设备数据库所存储关系型日志数据，则是通过 Sqoop 组建进行数据导入，进入数据采集平台的 Hive 中。工作原理如图 2 所示。

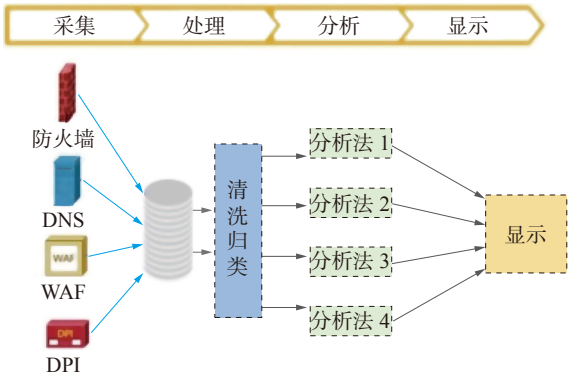

图2　平台工作原理

所采集到的日志数据，利用数据采集平台进行数据的解析，并进行存储。为了方便对数据采集平台中日志的有效分析及利用，开发日志数据采集平台的数据分析平台功能。

1.5　数据采集平台

数据采集平台部署了 Hadoop 集群的 Flume、Sqoop、HDFS、Hive、Spark 等组件，如图 3 所示，主要功能是采集、解析以及存储 IPv6 网络日志数据。

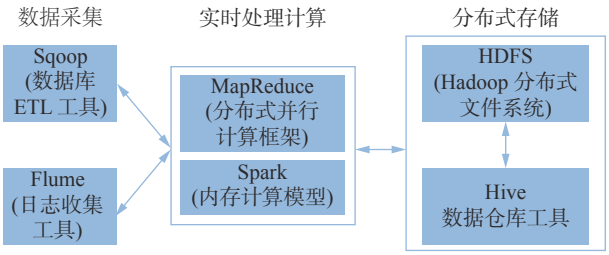

图3　CDH-Hadoop 集群各组件运行情况

1.5.1　数据采集

1) 原始日志数据采集

Flume 组件是 Cloudera 提供的日志收集系统，支持在日志系统中定制各类数据发送方，用于收集数据，主要实时采集非关系型的原始日志数据。

IPv6 网络日志采集平台，原始日志数据采集部分就是通过 Flume 组件对 IPv6 的 DNS 日志、防火墙的日志进行采集。图 4 为 Flume 采集过程。

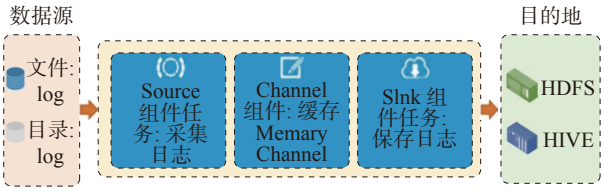

图4　Flume 采集过程

以下是采集过程中所涉及到的配置，业务信息已进行脱敏处理。

① 采集防火墙日志的 Flume 组件配置

```
# r4 config
a1.sources.r4.type = syslogudp
a1.sources.r4.channels = c4
a1.sources.r4.host = 0.0.0.0
a1.sources.r4.port = 600
a1.sources.r4.keepFields = true
# c4 config
a1.channels.c4.type = memory
a1.channels.c4.capacity = 10000
a1.channels.c4.transactionCapacity = 1000
# k4 config
a1.sinks.k4.type = FlumeHDFSPathSink
a1.sinks.k4.path = hdfs://xxx.xxx.xxx.xxx:xxxx/user/hdfs/firewall_数据中心_9010/
a1.sinks.k4.rollTimeType = minute
a1.sinks.k4.rollTimePeriod = 10
a1.sinks.k4.batchSize = 100
a1.sinks.k4.isCompress = 0
a1.sinks.k4.channel = c4
```

② 采集 DNSv6 日志的 Flume 组件配置

```
# r8 DNSv6 日志
a1.sources.r8.type = syslogudp
a1.sources.r8.channels = c8
a1.sources.r8.host = 0.0.0.0
a1.sources.r8.port = 621
# c8 config
a1.channels.c8.type = memory
a1.channels.c8.capacity = 10000
a1.channels.c8.transactionCapacity = 1000
# k8 config
a1.sinks.k8.type = FlumeHDFSPathSink
a1.sinks.k8.path = hdfs://xxx.xxx.xxx.xxx:xxxx/user/hdfs/DNSV6/
a1.sinks.k8.rollTimeType = minute
a1.sinks.k8.rollTimePeriod = 5
a1.sinks.k8.batchSize = 100
a1.sinks.k8.isCompress = 0
a1.sinks.k8.channel = c8
```

2) 关系数据库日志数据采集

Sqoop 组件是一个数据传输工具，专为大数据批量传输设计，主要用于分布式存储系统与关系型数据库间的数据传递[8]，在数据源与 Hadoop 之间进行数据转移时，可以保障数据处理的安全。

由于学校的 WAF 以及 DPI 设备均是采用 MySQL 数据库来存储相关日志，因此数据采集平台通过 Sqoop 组件对 WAF 日志数据、DPI 日志数据进行采集，导入到 Hadoop 的 HDFS、HIVE 等数据存储系统。图 5 为 Sqoop 导入过程。

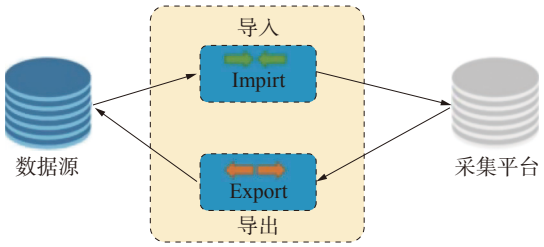

图5　Sqoop 导入过程

1.5.2　数据存储

HDFS 分布式文件系统提供海量数据分布式存储，由主节点和多个子节点组成，数据存储具有高可靠性和完整性。HDFS 基于移动计算的批处理技术，数据的访问采取一次写入多次读取的模式，保证数据的一致性和高吞吐率[5]。

IPv6 网络日志采集平台通过 Flume 组件所采集到的日志数据会直接存储在 HDFS 分布式文件系统的不同目录当中，利用开发的数据分析平台对 HDFS 目录进行查询分析，如图 6 所示。

<190>Nov 11 02:29:18 2021 SrcZoneName(1025)=cernet_ipv6;DstZoneName(1035)=Trust;Type(1067)=ACL6;RuleID(1078)=6111;Protocol(1001)=TCP;Application(1002)=http;
SrcIPv6Addr(1036)=2409:8947:5e5b:2b6:5834:5163:6bb1:369a;SrcPort(1004)=48484;
DstIPv6Addr(1037)=2001:da8:b807:0:202:118:176:28;DstPort(1008)=80;MatchCount(1069)=1;Event(1048)=Deny;
<190>Nov 11 00:53:54 2021 SrcZoneName(1025)=cernet_ipv6;DstZoneName(1035)=Trust;Type(1067)=ACL6;RuleID(1078)=6105;Protocol(1001)=UDP;Application(1002)=dns;
SrcIPv6Addr(1036)=240e:4a:4300:2::4;SrcPort(1004)=9712;
DstIPv6Addr(1037)=2001:da8:b807:0:218:7:43:4;DstPort(1008)=53;MatchCount(1069)=1;Event(1048)=Permit;
<190>Nov 11 04:39:27 2021 SrcZoneName(1025)=cernet_ipv6;DstZoneName(1035)=Trust;Type(1067)=ACL6;RuleID(1078)=6105;Protocol(1001)=UDP;Application(1002)=dns;
SrcIPv6Addr(1036)=2409:803c:2001:8003::2;SrcPort(1004)=38040;
DstIPv6Addr(1037)=2001:da8:b807:0:218:7:43:4;DstPort(1008)=53;MatchCount(1069)=1;Event(1048)=Permit;
<190>Nov 11 05:02:17 2021 SrcZoneName(1025)=cernet_ipv6;DstZoneName(1035)=Trust;Type(1067)=ACL6;RuleID(1078)=6105;Protocol(1001)=UDP;Application(1002)=dns;
SrcIPv6Addr(1036)=2620:171:fc:f0::240;SrcPort(1004)=8209;
DstIPv6Addr(1037)=2001:da8:b807:0:218:7:43:4;DstPort(1008)=53;MatchCount(1069)=1;Event(1048)=Permit;
<190>Nov 11 06:07:22 2021 SrcZoneName(1025)=cernet_ipv6;DstZoneName(1035)=Trust;Type(1067)=ACL6;RuleID(1078)=6108;Protocol(1001)=TCP;Application(1002)=http;
SrcIPv6Addr(1036)=2409:8978:35:788:2:1:3327:ba91;SrcPort(1004)=56646;
DstIPv6Addr(1037)=2001:da8:b807:0:202:118:176:8;DstPort(1008)=80;MatchCount(1069)=1;Event(1048)=Permit;
<190>Nov 11 06:12:22 2021 SrcZoneName(1025)=cernet_ipv6;DstZoneName(1035)=Trust;Type(1067)=ACL6;RuleID(1078)=6108;Protocol(1001)=TCP;Application(1002)=http;
SrcIPv6Addr(1036)=2409:8978:35:788:2:1:3327:ba91;SrcPort(1004)=56646;
DstIPv6Addr(1037)=2001:da8:b807:0:202:118:176:8;DstPort(1008)=80;MatchCount(1069)=2;Event(1048)=Permit;

图 6　查看 HDFS 上存储的原始日志格式

1.5.3　数据仓库

Hive 是 Hadoop 的数据仓库工具，将结构化的数据文件映射为一张数据库表，使用 HDFS 作为存储底层，Map Reduce 作为执行底层，并提供简单的查询功能[8]。

数据采集平台经过模板对原始日志的处理加工后的数据最终会通过 Hive 进入数据仓库，以满足 Spark SQL 的计算需求，如图 7 所示。

WAF0HW12C807	2022-04-14 09:56:50	111.7.106.100	7008	218.7.43.8	80
WAF0HW12C807	2022-04-14 09:58:40	2001:da8:b807:c1da:9481:6a2d:62ed:e8e3	1967	2001:da8:b807:0:202:118:176:28	80
WAF0HW12C807	2022-04-14 09:57:39	111.7.106.100	12182	218.7.43.8	80
WAF0HW12C807	2022-04-14 10:10:38	2409:891b:218:6d2:ed5c:b793:6719:eade	60248	2001:da8:b807:0:202:118:176:28	80
WAF0HW12C807	2022-04-14 10:12:56	10.42.97.54	42392	218.7.43.28	80
WAF0HW12C807	2022-04-14 09:55:36	2409:891a:21c:e48:cdb0:a4c2:da3f:2dd9	39862	2001:da8:b807:0:202:118:176:28	80
WAF0HW12C807	2022-04-14 10:08:27	121.227.152.92	65132	202.118.176.33	80
WAF0HW12C807	2022-04-14 09:51:59	2001:da8:b807:c28f:d483:e3c9:56d:fb0d	55578	2001:da8:b807:0:202:118:176:28	80
WAF0HW12C807	2022-04-14 09:57:37	111.7.106.100	11943	218.7.43.8	80
WAF0HW12C807	2022-04-14 10:07:04	10.40.122.56	49681	218.7.43.28	80

图7　查看经过解析处理后的日志格式

1.6　数据分析平台

数据分析平台采用 B/S 架构设计，业务层采用 Python Django 框架开发，平台存储采用 MySQL 数据库，表现层使用包括 jQuery、Bootstrap、ECharts 等前端技术实现交互，底层通过 Java 开放的 JAR 与 Hadoop 的 Spark 相关组件进行调度。Spark 是基于 Map Reduce 实现的分布式计算框架，将任务的中间输出及结果保存在内存中，能更好地适用于数据挖掘与机器学习等需要迭代的环境[8]。

数据分析平台总体功能划分为数据解析功能、数据查询功能以及数据分析功能。部署环境在 Linux 操作系统（Ubuntu 14.0.4+）Docker 模式下一键式安装和管理。

1.6.1　数据解析

数据解析功能主要是采用图形化界面的操作方式，通过调用已建立的 Spark 离线任务、Spark Streaming 流式任务将采集防火墙、DNS 的 IPv6 非结构化网络日志数据解析为结构化数据，存储到 Hive 数据库中，如图 8 所示。

图8　日志解析模块

1.6.2　数据查询

1) 防火墙 IPv6 流量访问明细查询

可以实时查询并展示防火墙中保存的 IPv6 访问数据，包括源地址、目的地址和目的端口，对热点业务进行分析，如图 9 所示。

2) WAF IPv6 流量过滤明细查询

可以实时查询并展示 WAF 中保存的 IPv6 攻击数据，包括源地址与端口、目的地址与端口，对攻击进行分析溯源。如图 10 所示。

1.6.3　数据分析

采用图形化界面操作，通过调用 Spark SQL 对异构数据源（Hive、MySQL）进行数据抽取转换、处理以及分析计算，并将最终结果输出到结构化数据库 Hive 中，如图 11 所示。另外表 1 为经过计算后的样例数据。

四层协议名称(protocol)	应用名称(Application)	源IP地址(SrcIPv6Addr)	源端口号(SrcPort)	目的IP地址(DstIPv6Addr)
UDP	syslog	2001:250:7801:a0::ac12:202	39799	2001:da8:b807:0:202:118:177:170
UDP	syslog	2001:250:7801:a0::ac12:202	39799	2001:da8:b807:0:202:118:177:170
UDP	syslog	2001:250:7801:a0::ac12:204	54240	2001:da8:b807:0:202:118:177:170
UDP	syslog	2001:250:7801:a0::ac12:204	54240	2001:da8:b807:0:202:118:177:170
UDP	syslog	2001:250:7801:a0::ac12:202	39799	2001:da8:b807:0:202:118:177:170
UDP	syslog	2001:250:7801:a0::ac12:204	54240	2001:da8:b807:0:202:118:177:170
UDP	syslog	2001:250:7801:a0::ac12:204	54240	2001:da8:b807:0:202:118:177:170
UDP	syslog	2001:250:7801:a0::ac12:204	54240	2001:da8:b807:0:202:118:177:170
UDP	syslog	2001:250:7801:a0::ac12:202	39799	2001:da8:b807:0:202:118:177:170
UDP	syslog	2001:250:7801:a0::ac12:204	54240	2001:da8:b807:0:202:118:177:170

图9　防火墙 IPv6 流量访问明细查询

2022-04-01 00:05:57	2001:da8:b807:d90e:502d:21e5:f6f2:65b	41188	2001:da8:b807:0:202:118:176:28	80
2022-04-01 00:05:57	2001:da8:b807:d90e:502d:21e5:f6f2:65b	41218	2001:da8:b807:0:202:118:176:28	80
2022-04-01 00:05:57	2001:da8:b807:d90e:502d:21e5:f6f2:65b	41184	2001:da8:b807:0:202:118:176:28	80
2022-03-31 23:43:11	2001:250:7801:10:f144:b3ed:2b59:59d9	54737	2001:da8:b807:0:202:118:176:28	80
2022-03-31 23:18:27	2001:da8:b807:d90e:502d:21e5:f6f2:65b	38322	2001:da8:b807:0:202:118:176:28	80
2022-03-31 23:18:27	2001:da8:b807:d90e:502d:21e5:f6f2:65b	38326	2001:da8:b807:0:202:118:176:28	80
2022-03-31 23:18:27	2001:da8:b807:d90e:502d:21e5:f6f2:65b	38326	2001:da8:b807:0:202:118:176:28	80
2022-03-31 23:18:27	2001:da8:b807:d90e:502d:21e5:f6f2:65b	38322	2001:da8:b807:0:202:118:176:28	80

图10　WAF IPv6 流量访问明细查询

190	2022-04-14 09:40:10	M9006-IRF	%%10	FILTER	6	FILTER ZONE IPV6 EXECUTION	-Chassis=1-Slot=5.1	cernet ipv6	Trust
190	2022-04-14 09:55:10	M9006-IRF	%%10	FILTER	6	FILTER ZONE IPV6 EXECUTION	-Chassis=1-Slot=5.1	cernet ipv6	Trust
190	2022-04-14 10:07:43	M9006-IRF	%%10	FILTER	6	FILTER ZONE IPV6 EXECUTION	-Chassis=1-Slot=5.1	cernet ipv6	Trust
190	2022-04-14 09:17:44	M9006-IRF	%%10	FILTER	6	FILTER ZONE IPV6 EXECUTION	-Chassis=1-Slot=5.1	cernet ipv6	Trust
190	2022-04-14 10:00:10	M9006-IRF	%%10	FILTER	6	FILTER ZONE IPV6 EXECUTION	-Chassis=1-Slot=5.1	cernet ipv6	Trust
190	2022-04-14 10:02:44	M9006-IRF	%%10	FILTER	6	FILTER ZONE IPV6 EXECUTION	-Chassis=1-Slot=5.1	cernet ipv6	Trust
190	2022-04-14 09:22:44	M9006-IRF	%%10	FILTER	6	FILTER ZONE IPV6 EXECUTION	-Chassis=1-Slot=5.1	cernet ipv6	Trust
190	2022-04-14 09:52:44	M9006-IRF	%%10	FILTER	6	FILTER ZONE IPV6 EXECUTION	-Chassis=1-Slot=5.1	cernet ipv6	Trust
190	2022-04-14 09:25:10	M9006-IRF	%%10	FILTER	6	FILTER ZONE IPV6 EXECUTION	-Chassis=1-Slot=5.1	cernet ipv6	Trust
190	2022-04-14 09:32:44	M9006-IRF	%%10	FILTER	6	FILTER ZONE IPV6 EXECUTION	-Chassis=1-Slot=5.1	cernet ipv6	Trust

图11　通过 Spark SQL 的技术结果预览

表1　经过计算后的主要数据样例

日志产生时间	4层协议名称	应用名称	源IP地址源	端口号	目的IP地址	目的端口号
2021-01-15 07:59:14	UDP	dns	2409:8038:2000:206d::1	53668	2001:da8:b807:0:218:7:43:4	53
2021-01-15 07:59:31	TCP	general_tcp	240e:f7:4f01:c::3	43257	2001:da8:b807:0:202:118:176:3f05	9944
2021-01-15 07:59:57	TCP	http	240e:468:118:1c98:119d:2377:3bb3:7ad0	43152	2001:da8:b807:0:202:118:176:8	80
2021-01-15 07:59:59	UDP	dns	240e:3:3800:ff22::3	55930	2001:da8:b807:0:202:118:176:2	53
2021-01-15 08:00:23	UDP	dns	2408:8001:2007:1::3e	52238	2001:da8:b807:0:218:7:43:4	53

2　可视化监控分析

对 IPv6 网络流量日志解析获取时间、协议、应用、源地址及端口号和目的地址及端口号等相关数据，并进行联合数据分析。通过分析结果的可视化展示，可以效率化分析研判网络安全威胁[9]，使高校数据中心运维人员通过更直观地发现 IPv6 流量传输的规律、分析某段时间内网络的负载情况、观察上网用户的行为趋势、发现网络中的异常甚至防范网络攻击等。从而方便运维人员对 IPv6 网络的管理和运维[10]。

本文结合数据可视化监控需求和数据变化影响关系，利用 JAVA 开发语言中的可视化展现插件，与日志分析平台进行融合，开发设计可视化监控分析功能，形成具有编辑能力的数据展示监控平台。

具体功能如下：

1) 高校数据中心 24 h 内 IPv6 源地址访问分析

可以展示访问数据中心的来源地址数据，统计 IPv6 地址来源详情，如图 12。

图12　数据中心 24 h 内 IPv6 源地址访问分析

2) 高校数据中心 24 h 内 IPv6 目的地址访问分析

可以展示数据中心内部地址被访问详情，统计内部 IPv6 业务访问情况，如图 13。

图 13　数据中心 24 h 内 IPv6 目的地址访问分析

3) 高校数据中心 24 h 内 IPv6 端口访问分析　　计未知端口被扫描和攻击情况，如图 14。可以展示数据中心内部端口被访问详情，统

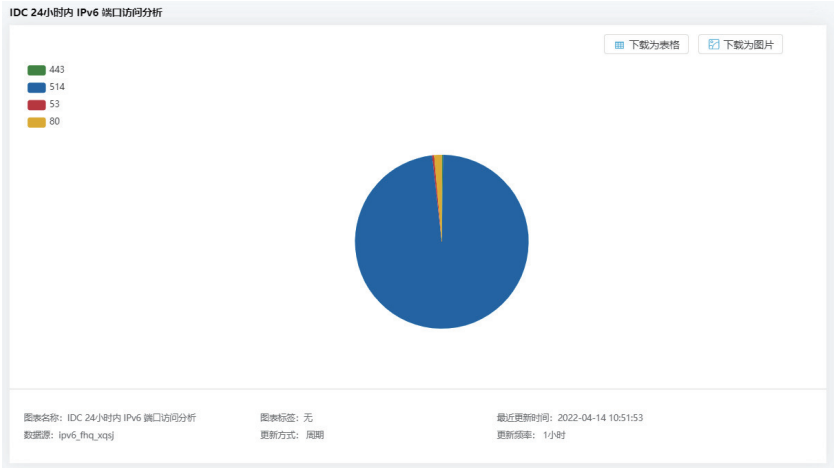

图14　数据中心 24 h 内 IPv6 端口访问分析

4) 高校数据中心 24 h 内 IPv6 地址攻击详情　　时对攻击主机进行加固和安全防护，如图 15。可以展示数据中心内部地址被攻击详情，及

图15　数据中心 24 h 内 IPv6 地址攻击详情

5) 高校数据中心 24 h 内 IPv6 源地址攻击详情　　对攻击来源进行溯源和封堵，如图 16。可以展示 IPv6 地址攻击数据中心内部详情，

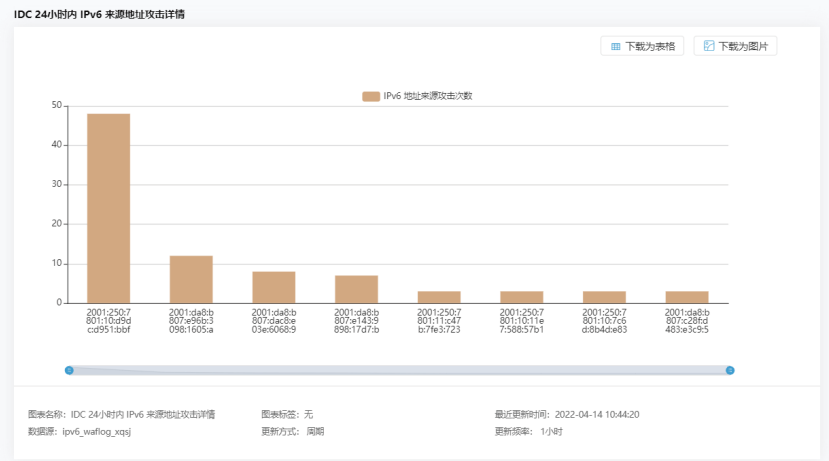

图 16　数据中心 24 h 内 IPv6 源地址攻击详情

3　结论

随着 IPv6 协议不断完善和国内外对 IPv6 网络及应用的加快推进和部署, 对于走在科技前沿的高等院校来说, 将迎来 IPv6 网络及应用的快速建设和全面实施。由于当下主流的网络与安全的软硬件设备都对 IPv6 协议有较好支持, 设计实现的数据中心 IPv6 网络日志采集平台, 能够实时统一采集存储各类软硬件 IPv6 日志信息, 快速高效查询分析日志信息, 既减轻高校数据中心 IPv6 网络管理运维的工作体量与复杂度, 又可通过对日志数据的关联分析, 为 IPv6 网络攻击的检测预警工作提供依据。未来可采集更多类型的数据中心网络及应用的 IPv6 日志, 继续对日志数据进行深度挖掘与分析, 为高校数据中心 IPv6 网络的管理运维和安全监测提供全面详细的数据支撑。

参考文献:

[1] 国家 IPV6 发展监测平台 [EB/OL]. (2021−12−08). https://www.china-IPv6.cn.

[2] 王维. IPv6 网络安全威胁的分析与研究 [J]. 中国管理信息化, 2021, 24(17): 178−180.

[3] 秦道祥, 路阳, 张荠月, 等. 基于 Spark 技术的日志分析平台设计与应用 [J]. 中国教育信息化, 2021(19): 50−54.

[4] 陈建锐. 基于协议分析的 IPV6 入侵检测系统研究 [J]. 计算机与数字工程, 2011, 39(9): 98−100, 135.

[5] 尹中江. 基于 Hadoop 的大数据基础平台搭建与西藏农业应用构想 [J]. 西藏农业科技, 2019, 41(3): 73−78.

[6] 莹, 周承敏, 张晓桐. 基于大数据多源安全日志的设计与实现 [J]. 网络空间安全, 2021, 12(S2): 28−33.

[7] 游会迪, 张振友. 基于 Hadoop 大数据平台的搭建及其测试研究 [J]. 电脑知识与技术, 2017, 13(19): 211−213.

[8] 马浩, 武超飞, 王立斌, 等. 基于 Hadoop CDH 的用电信息大数据平台的研究 [J]. 河北电力技术, 2018, 37(2): 33−34, 59.

[9] 王宇, 温占考, 王卫东, 等. 校园网络 IPv4/IPv6 威胁监测与处置体系的规划与实践 [J]. 深圳大学学报 (理工版), 2020, 37(S1): 55−59.

[10] 吕悦. 基于 IPv6 的网络日志数据分析系统 [J]. 信息技术与信息化, 2021(8): 50−53.

本文引用格式:

吴宇平, 邢丽刃, 秦绪杰. 数据中心 IPv6 网络日志采集平台的设计与实现 [J]. 应用科技, 2022, 49(3): 76–83.

WU Yuping , XING Liren, QIN Xujie. Design and implementation of IPv6 network log collection platform in data center[J]. Applied science and technology, 2022, 49(3): 76–83.